

High Availability for the Cisco Catalyst 6500 Series Switches

Overview

Cisco Catalyst® 6500 Series multilayer switches have become an essential component of a sound network design in today’s enterprise and service provider environments. Having such a critical role, the Cisco Catalyst 6500 Series must provide a reliable switching platform, and offer high performance and intelligent network services. The high availability of the Cisco Catalyst 6500 Series even has the capability to maintain an IP phone call during supervisor engine failover. This paper discusses how the Cisco Catalyst 6500 Series provides high system availability through hardware and software redundancy features, and focuses specifically on the following three areas:

Fabric redundancy of the Switch Fabric Module (SFM)

Supervisor engine redundancy with the Cisco Catalyst Operating System (Catalyst OS), High Availability feature, which includes the stateful protocol redundancy and image versioning functions

- Multilayer Switch Feature Card (MSFC) Cisco IOS® Software redundancy features—Dual Router Mode (DRM), Configuration-Synchronization (config-sync), and Single Router Mode (SRM).

This paper is based on the hybrid software model for the Cisco Catalyst 6500 Series (Cisco Catalyst OS on the supervisor engine, Cisco IOS Software on the MSFC) and not on the Cisco IOS Software model (native Cisco IOS Software). All feature set references will be specifically described as a Cisco Catalyst OS feature on a supervisor engine or a Cisco IOS Software feature on an MSFC. The Cisco Catalyst OS High Availability feature was first introduced in the Cisco Catalyst OS 5.4 release and is available for both Cisco Catalyst Supervisor Engine 1A and Catalyst Supervisor Engine 2. Support for DRM began in Cisco IOS Software Release 12.0(7)XE1. The MSFC config-sync redundancy feature for DRM is supported in Cisco IOS Software Release 12.1(3a)E4 for both the MSFC and MSFC2. The MSFC SRM feature was first supported with Cisco Catalyst OS 6.3.1 and Cisco IOS Software Release 12.1(8)E2 for the MSFC2.

This paper is the second version of the original that was written in September 2000. This version includes some updated sections for more precise understanding and a discussion of SRM.

Although component-level redundancy is very important, a high-availability network design relies on the proper combination of individual system redundancy and overall network

Figure 1
 The Cisco Catalyst 6500 Series WS-6503, WS-C6506, WS-C6509, WS-C6509-NEBS, and WS-C6513





redundancy. For more detail on high-availability network designs, refer to the white paper, *Gigabit Campus Design*, at: http://www.cisco.com/warp/public/cc/so/neso/Inso/cps/camp_wp.htm.

Switch Fabric Module Redundancy

Since its introduction, the Cisco Catalyst 6500 Series has been built on a single 32-Gbps bus switching architecture that provides the data path for all packets through the system. The Cisco Catalyst 6500 Series includes a 256-Gbps crossbar switching fabric (the SFM for higher bandwidth capacities and 30+ Mpps of forwarding performance). The SFM is supported in the Cisco Catalyst 6506 and the Cisco Catalyst 6509 chassis. The SFM2 is essentially the same fabric but designed to work in all the Cisco Catalyst 6506, 6509, and 6513.

Switching Fabric Failover

The SFM also provides another level of hardware redundancy to the system. The single fabric channel versions of the fabric-enabled line cards provide a connection to both the switching fabric and the existing system bus backplane. This allows the Cisco Catalyst 6500 Series to use the SFM as the primary data path between fabric-enabled line cards. In the event that an SFM fails, the system will fail over to the 32-Gbps bus to ensure that packet switching continues (albeit at the bus capacities of 15 Mpps throughput and 32-Gbps bandwidth) and the network remains online. Additionally, a Cisco Catalyst 6500 Series can be configured with dual SFMs (in slots 5 and 6 of a Catalyst 6506 or Catalyst 6509 or in slots 7 and 8 of a Catalyst 6513), which provide a third level of fabric redundancy. In this configuration, a failure on the primary fabric module would result in a switchover to the secondary fabric module for continued operation at 30 Mpps. Also, in the event of further fabric module failures, the ability to switch over yet again to the system bus would still be available.

Switching Fabric Operation

Different combinations of SFMs, fabric-enabled line cards, and classic line cards in a chassis affect the internal switching operation, which in turn affects the failover characteristics. This is an important point to understand as fabric-to-fabric or fabric-to-bus failover scenarios are discussed. When an SFM is installed in a system of only fabric-enabled line cards, the switching operation is called compact mode. This allows for 32-byte compacted headers (not the entire packet) to be sent across the bus to the supervisor engine for each forwarding decision. The increase in efficiency for this operation allows for inherent system performance capable of 30 Mpps. The data path for fabric-enabled cards is via the SFM.

If a classic line card is installed in a system with an SFM, the header format on the bus must be compatible with all the line cards in the system. Because classic line cards do not support compact mode, the fabric-enabled line cards will change their switching modes to truncated mode. Truncated mode allows the fabric-enabled line card to send packets in a 64-byte header-only format that the classic line cards can understand. It is very important to note that the truncated mode still uses the SFM as the data path between fabric-enabled line cards. Although the maximum centralized forwarding performance is 15 Mpps in a system of classic and fabric cards, the switch fabric is still used to provide higher bandwidth to the system. If fabric-enabled line cards are installed in a system with no SFM, they will operate in flow-through mode even when classic cards are present. This mode essentially programs the line card to operate in a classic mode whereby the entire packet is sent across the system bus for a forwarding decision. A system in flow-through mode is capable of switching 15 Mpps and the data path is via the 32-Gbps bus.



The changes to the switching mode are done automatically, depending on the hardware installed. No specific configuration is necessary on the SFM for typical operation. The current switching mode of the switch fabric module can be monitored through the Catalyst OS command-line interface (CLI) using the **show fabric channel switchmode** command. Example 1 shows a completely fabric-enabled system (all compact mode) and Example 2 shows classic and fabric-enabled line cards in an SFM system (flow-through and truncated mode).

Example 1: Fabric-Enabled System

The following output is from a configuration with dual supervisor engines, one SFM, and a fabric-enabled line card in slot 3.

```
Sup2-A> (enable) show fabric channel switchmode
Module Num Fab Chan Fab Chan Switch Mode Channel Status
-----
 1      1 0, 0 compact ok
 2      1 0, 1 compact ok
 3      1 0, 2 compact ok
 5     18 0, 0 n/a ok
 5     18 1, 1 n/a ok
 5     18 2, 2 n/a ok
 5     18 3, 3 n/a unknown
 5     18 4, 4 n/a unknown
 5     18 5, 5 n/a unknown
 5     18 6, 6 n/a unknown
 5     18 7, 7 n/a unknown
 5     18 8, 8 n/a unknown
 5     18 9, 9 n/a unknown
 5     18 10, 10 n/a unknown
 5     18 11, 11 n/a unknown
 5     18 12, 12 n/a unknown
 5     18 13, 13 n/a unknown
 5     18 14, 14 n/a unknown
 5     18 15, 15 n/a unknown
 5     18 16, 16 n/a unknown
 5     18 17, 17 n/a unknown
15      0 n/a n/a n/a
16      0 n/a n/a n/a
```

CLI output description for **show fabric channel switchmode**:

Num Fab Chan—The number of fabric channels that the module is associated with.

Fab Chan—The first number is the fabric channel number that the module is associated with. The second number is the fabric channel number that the SFM is associated with.

Switch mode—Possible output is “flow through,” “truncated,” “compact.” Switch mode applies only to line cards with fabric and bus connections.

Channel status—Possible output is “ok,” “sync error,” “CRC error,” “heartbeat error,” “buffer error,” “timeout error,” or “unknown.” Channel status applies only to line cards with fabric and bus connections.



Example 2: Classic and Fabric-Enabled System

The following output is from a configuration with dual supervisor engines, one SFM, one classic line card in slot 3, and two fabric-enabled line cards in slots 7 and 9.

```
Sup-A> (enable) show fabric channel switchmode
Module Num Fab Chan Fab Chan Switch Mode Channel Status
-----
 1      1 0, 0 flow through ok
 2      1 0, 1 truncated ok
 3      0 n/a n/a n/a
 5     18 0, 0 n/a ok
 5     18 1, 1 n/a ok
 5     18 2, 2 n/a unknown
 5     18 3, 3 n/a unknown
 5     18 4, 4 n/a unknown
 5     18 5, 5 n/a unknown
 5     18 6, 6 n/a ok
 5     18 7, 7 n/a unknown
 5     18 8, 8 n/a ok
 5     18 9, 9 n/a unknown
 5     18 10, 10 n/a unknown
 5     18 11, 11 n/a unknown
 5     18 12, 12 n/a unknown
 5     18 13, 13 n/a unknown
 5     18 14, 14 n/a unknown
 5     18 15, 15 n/a unknown
 5     18 16, 16 n/a unknown
 5     18 17, 17 n/a unknown
 7      1 0, 6 truncated ok
 9      1 0, 8 truncated ok
15      0 n/a n/a n/a
16      0 n/a n/a n/a
```

Automatic switching mode changes allow the acceptance of classic or fabric-enabled cards into a system with no manual configuration change. As previously stated, there is a performance versus interoperability tradeoff when installing classic line cards into a fabric-enabled system. Because many network environments hold performance in higher regard, a fabric-enabled system can be configured to reject classic cards (for example, not support flow-through mode). By issuing the **set system crossbar-fallback none** command, the system will not start classic line cards installed in the chassis, thereby running in compact switching mode (30 Mpps) only.

```
Sup-A> (enable) set system crossbar-fallback none
```

The default for the crossbar-fallback is bus mode. To determine the current system state, the **show system crossbar-fallback** command is available.

```
Sup-A> (enable) show system crossbar-fallback
Cross-fallback: bus-mode
```

In summary, the SFM can be redundantly configured in a chassis to provide fabric-to-fabric and fabric-to-bus failover. A system configured with dual SFMs can use the standby SFM for failover. Additionally, single SFM systems with fabric-enabled line cards can fail over to the 32-Gbps bus for continuous operation. In both of these scenarios, recovery and return to normal operation occur in less than three seconds. This quick recovery time allows for a switching mode change and a synchronization process that must take place between each line card, supervisor engine, and the SFM fabric channels in these scenarios. The capability to configure redundant SFMs provides up to three levels of backplane redundancy, helping to enable continuous operations with minimal impact to network availability in the event of a hardware failure.



Redundant Supervisor Engines

As previously mentioned, the High Availability feature on the Cisco Catalyst 6500 Series provides low-impact, stateful switchover between redundant supervisor engines. This feature was first available in Cisco Catalyst OS Software Version 5.4.

Supervisor Engine Switchover

Dual supervisor engines provide hardware redundancy for the forwarding intelligence of the Cisco Catalyst 6500 Series. The Cisco Catalyst 6500 Series can support up to two supervisor engines in slots 1 and 2 only. One is the active supervisor engine and the other is the standby supervisor engine. The active supervisor engine is the first one to go online. This can be confirmed by the “Active” LED on the supervisor engine or by typing the **show module** command from the console. Both supervisor engines *must* be the same hardware models. This means that if a Policy Feature Card (PFC) and a MSFC are on a Supervisor 1A in slot 1, then a PFC and MSFC must be also on a Supervisor Engine 1A in slot 2, or if a Supervisor Engine 2 is in slot 1, a Supervisor Engine 2 must also be in slot 2. Supervisor engines 1A and 2 can be used in the Cisco Catalyst 6000 and 6500 series. If an active supervisor is taken offline or fails, the standby supervisor takes control of the system.

The two supervisor engines in a redundant supervisor configuration have different responsibilities. The active supervisor engine is responsible for controlling the system bus and all line cards. All protocols are running on the active supervisor engine and it performs all packet forwarding. The standby supervisor engine does not communicate with the line cards. It receives packets from the network and populates its forwarding tables with this information but does not participate in any packet forwarding. The relevant protocols on the system are initialized, but not active, on the standby supervisor engine. The Cisco Catalyst 6500 Series supervisor engines are hot swappable and the standby supervisor engine can be installed in an active system without affecting network operation. Also important to note is that redundant supervisor engines do not perform load sharing. The active supervisor engine is providing the entire packet forwarding intelligence for the system (N+1 redundancy). If the active supervisor engine fails, the standby supervisor engine can maintain the same system load.

The standby supervisor engine polls the active supervisor engine via the Ethernet out-of-band channel (EOBC) every 5–10 milliseconds to monitor the online status of the active supervisor engine. The active supervisor engine may go offline for a variety of reasons such as hardware failures, system overload conditions, memory corruption issues, removal from chassis, or being reset by the operator. The standby supervisor engine detects this type of failure and becomes the new active supervisor engine. The Cisco Catalyst OS software on the supervisor engine is responsible for restoring the protocols, line cards, and forwarding engines to normal operation. This restoration takes place via a fast switchover or a high-availability switchover.

Supervisor Fast Switchover

Because the Cisco Catalyst OS High Availability feature is disabled by default, the alternative is referred to as Fast Switchover. The Fast Switchover feature is the predecessor to the High Availability feature and as such is the supervisor switchover mechanism in place when high availability is disabled or not supported in the software version. This feature reduces the switchover time by skipping some events that would typically take place should a supervisor fail. Specifically, the fast switchover mechanism allows each line card to skip the respective software downloads and a portion of the diagnostics, which are normally a part of system re-initialization. The switchover still includes restarting all protocols at Layer 2 and above as well as resetting all ports. The resulting switchover performance with default settings will take approximately 28 seconds plus the time it takes for the protocols to restart. As an example, a switch with the default time values for the Spanning-Tree Protocol took approximately 58 seconds after the fast switchover to begin forwarding traffic again. However, the time to begin forwarding traffic after a fast switchover can be reduced by tuning the switch from the default settings. By enabling Portfast,



disabling port channels (PagP), and turning trunking off for ports to which workstations are directly attached, the fast switchover time can be reduced to approximately 10 seconds. In a live network environment, these switchover times present a major disruption to network operations.

Supervisor High Availability Feature

The High Availability software feature of Cisco Catalyst OS further enhances the Cisco Catalyst 6500 Series hardware redundancy by also providing protocol redundancy. This feature includes stateful protocol redundancy and image versioning. The High Availability feature must be enabled via the CLI for these features to operate.

```
Sup-A> (enable) set system highavailability enable  
System high availability enabled.
```

As a general practice with redundant supervisors, it is recommended that the High Availability feature be enabled for normal operation.

Supervisor Stateful Protocol Redundancy

The stateful supervisor switchover is when the switchover time from the active to the standby supervisor is reduced to less than three seconds. This reduced downtime is achieved by synchronizing many of the Layer 2, Layer 3, and Layer 4 protocols¹ between the active and standby supervisor engines and is called *maintaining protocol state*.

For stateful protocol redundancy between dual supervisor engines, a protocol state database is maintained on each supervisor engine for all protocols and features requiring high-availability support. Most of these protocols are only running on the active supervisor engine. In the event of a high-availability switchover, the new active supervisor engine can start the protocols from the updated database state, rather than the initialization state. This is how a redundant system can maintain stateful protocol redundancy and minimal network downtime when the active supervisor engine goes offline.

- **High Availability Supported Feature**—High availability is fully supported. The state of the feature is preserved between the active and standby supervisor engines in the protocol database.
- **High Availability Compatible Feature**—High availability is not supported for these features. The protocol database for these features is *not* synchronized between supervisor engines. The feature *can* be used if the High Availability feature is enabled. For example, if GARP Multicast Registration Protocol (GMRP) and high availability were both enabled and a high-availability supervisor engine failover took place, the GMRP protocol would be restarted from the initialization state (non-stateful). The stateful protocol redundancy is still in place for the supported features if a compatible feature is enabled.
- **High Availability Incompatible Feature**—High availability is not supported. The protocol database for these features is *not* synchronized between supervisor engines. The feature should *not* be enabled if the High Availability feature is enabled. These features are not supported with high availability enabled because incorrect behavior may result.

Important: Do not use these features if a high-availability system is required.

1. Layer 4 protocols include the Layer 4 information in extended IP access lists.



Table 1 shows protocols and features for high availability as of Cisco Catalyst OS version 7.5 that are supported, compatible, and incompatible.

Table 1 High Availability Feature Support

Supported Features	Compatible Features	Incompatible Features
COPS-DR and COPS-DR	ASLB	Dynamic VLANs
Dynamic Trunk Protocol	Cisco Discovery Protocol	Generic VLAN Registration Protocol (GVRP)
Cisco Express Forwarding and adjacency tables	GMRP	Protocol filtering
Private VLANs	Internet Group Management Protocol (IGMP) snooping	
Router access control lists (ACLs)	Remote Monitoring (RMON)	
Multilayer switching (MLS)	Resource Reservation Protocol (RSVP)	
Port Aggregation Protocol/Link Aggregation Protocol (PAgP/LACP)	Simple Network Management Protocol (SNMP)	
Quality-of-service (QoS) ACLs and policers	Telnet sessions	
Switched Port Analyzer (SPAN)	VTP pruning	
STP	Uplinkfast	
Trunking		
UniDirectional Link Detection (UDLD) protocol		
VLAN ACLs		
VLAN Trunking Protocol (VTP)		
Port Security		
802.1X		

For a current list of the features that are supported with the High Availability feature, see the “Configuring Redundancy” chapter of the Cisco *Catalyst 6500 Series Software User Guide* and the release notes.

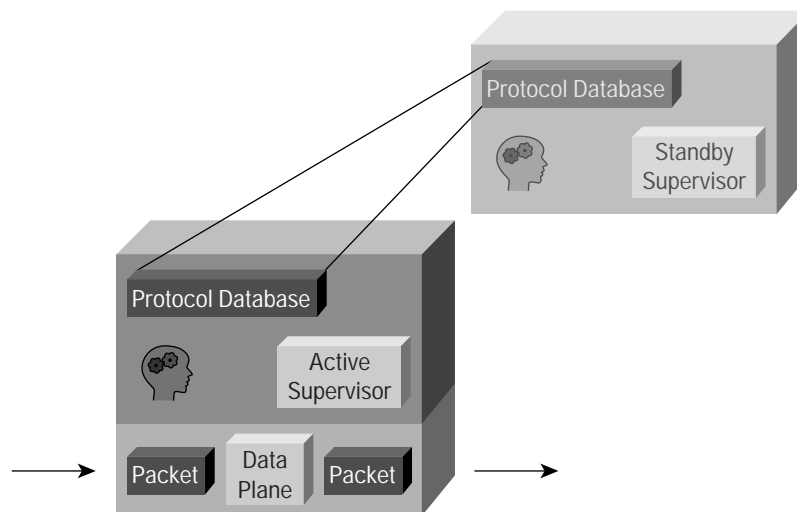
Many Layer 3 and Layer 4 protocols or features are programmed into the application-specific integrated circuits (ASICs) of the PFC or PFC2 on board the supervisor engine. Examples of these include access lists (router- and VLAN-based), forwarding tables (multilayer switching cache and Cisco Express Forwarding tables), and QoS settings. The protocols are maintained in the protocol database and will continue to be switched in hardware when a supervisor engine failover occurs. Some of these are dependent on a dual MSFC configuration, which is discussed later in this paper.

The protocol state database, depicted in Figure 2 below, is a repository of up-to-date protocol state information. It is generated by the active supervisor and stored by the standby supervisor. The database contains specific system information including module and port states, VLAN information, nonvolatile RAM (NVRAM) configurations, and various protocol-specific data.



Both supervisor engines run a synchronizing operation to allow for transfer of this data. When a database entry is updated on the active supervisor engine, the synchronizing operation places the update in a first-in, first-out (FIFO) queue. This queue is scheduled to empty periodically for transfer to the standby supervisor. The transfer is a background process and as such, the update interval varies depending on the number of other active processes in the system. The update interval ranges from one to five seconds with two seconds being an approximate average. The standby supervisor engine's synchronizing process receives these asynchronous updates and enters them into the protocol state database on the standby supervisor engine. When the system starts or when a second supervisor engine is hot-inserted, a global synchronization takes place between the protocol databases to ensure all protocol states are up to date.

Figure 2
Stateful Protocol Database Depiction



To summarize the stateful protocol feature, the high-availability switchover performance is more dependent on the status of the synchronization procedure than on the complexity of the configuration. After the system and protocols reach a stable operating point, the protocol state databases on each of the supervisor engines will have a fairly similar status (depending on updates in the queue). The determining factor for switchover performance is the number of updates in the queue that have not been completed. The resulting high-availability supervisor engine switchover performance is less than three seconds.

Supervisor Engine Software Image Upgrades

In a redundant supervisor engine configuration, Cisco Catalyst OS images must be properly managed to ensure high availability of the system. The following section describes some options for managing Cisco Catalyst OS images.

Supervisor Engine Image Synchronization

By default on the Cisco Catalyst 6500 Series, the Cisco Catalyst OS software images on the active and the standby supervisor engines must be the same. This allows the system to maintain a stable operating environment by ensuring that the supervisor engine switchover occurs with the same software features and revisions on the new active supervisor engine as on the previously active supervisor engine. If the two images are not the same version during system bootup, the active supervisor engine downloads its current boot image to the standby supervisor engine. The NVRAM configuration of the active supervisor engine is also synchronized between the supervisor engines.



The image synchronization feature of Cisco Catalyst OS provides software consistency between supervisor engines, but it does not allow for software upgrades without taking the system offline for an extended period of time. To perform the upgrade, the active supervisor engine requires a reset to load the new version of software. It then synchronizes the software images to the standby supervisor engine. Typically, this must be performed during a scheduled downtime or maintenance window because the entire system needs to be warm booted. Also note that the Cisco IOS Software on the MSFC is not a part of this synchronization process.

Supervisor Engine Versioning Feature

Versioning is the second portion of the Cisco Catalyst OS High Availability feature and is dependent on having the High Availability feature enabled in a dual supervisor engine configuration. As such, it allows different but compatible images to be running on the active and standby supervisor engines, which disables the default supervisor image synchronization process. The application of this feature is to allow a software upgrade in real time by using the supervisor switchover of the High Availability feature. This allows not only the upgrading of Cisco Catalyst OS software without rebooting the device, but also the ability to maintain a previously used and tested version of the Cisco Catalyst OS on the standby supervisor engine as a fallback plan if the software upgrade fails. There is no restriction on the image version that either supervisor engine can run, so upgrading or downgrading of the Catalyst OS images is possible.

If two different image versions are running, the system will determine if they are compatible. The active and standby supervisor engines exchange image version information to determine if the two software images are compatible. Image versions are defined as one of three options: compatible, incompatible, or upgradable. Compatible versions imply that stateful protocol redundancy can be supported between the different images. All configuration settings made to the NVRAM on the active supervisor engine can be sent to the standby supervisor. Two Cisco Catalyst OS versions are incompatible if synchronizing the protocol state databases between the two versions is not possible. If two software images are incompatible, the software upgrade process will affect the system operation (that is, be greater than the one- to three-second switchover time of a high-availability switchover) and no NVRAM configuration changes will be synchronized between supervisor engines. A special case of incompatible versions is referred to as upgradable. In this scenario, the high-availability supervisor switchover is not available, but configuration changes to the NVRAM on the active supervisor engine can be synchronized to the standby supervisor engine. This is a special case because it allows two different software versions to run with synchronized configurations but without the capability for a failover.

If the Cisco Catalyst OS software images are not compatible, the high-availability switchover will not be possible. The operational status output from the command **show system highavailability** should be monitored to determine the high-availability compatibility of two Cisco Catalyst OS images. The operational status can either be ON or OFF, with some system specific status messages. The following output shows that high availability is enabled and that the Cisco Catalyst OS versions are high-availability-compatible (Op-status: ON).

```
Sup-A> (enable) show system highavailability  
Highavailability: enabled  
Highavailability versioning: disabled  
Highavailability Operational-status: ON
```

As a general practice, it is recommended that high-availability versioning be enabled only when upgrading the Cisco Catalyst OS software. The traditional image synchronization process (high-availability versioning disabled) should be implemented for normal operating conditions. Generally speaking, high-availability compatible images are only available between maintenance releases of the Cisco Catalyst OS software. A maintenance release is a new version of software with incremental feature upgrades and bug fixes such as upgrading from version 5.5.1 to version 5.5.2. Major releases will not be high-availability compatible. The release notes include a high-availability compatibility listing at the following URL:



http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/relnotes/ol_1982.htm - xtocid8

Cisco Catalyst OS Image Upgrade Procedure

Based on the discussion above, the following procedure is recommended for performing a software upgrade with the minimal downtime associated with the High Availability feature. High-availability compatibility between the images will be determined in the middle of the procedure. The MSFC is affected by this procedure, but is discussed in the section “MSFC High Availability Features.”

In this example, the supervisor engine in slot 1 (Sup-A) will begin in active mode and the supervisor engine in slot 2 (Sup-B) will begin in standby mode. It is recommended that a console connection be available for both supervisors for this procedure.

1. Disable the High Availability feature on the active supervisor engine.

```
Sup-A> (enable) set system highavailability disable
```

This feature is disabled by default.

2. Load the new Cisco Catalyst OS software image into the boot Flash of the active supervisor engine (via slot 0, Trivial File Transfer Protocol [TFTP], etc.).

```
Sup-A> (enable) copy slot0:cat6000-sup2k8.7-2-2.bin bootflash:cat6000-sup2k8.7-2-2.bin
```

3. Verify that the new image was loaded successfully into the boot Flash of the active supervisor engine.

```
Sup-A> (enable) dir bootflash:
```

4. Clear the current boot variable.

```
Sup-A> (enable) clear boot system all
```

5. Set the boot variable on the active supervisor engine to the new Cisco Catalyst OS software image.

6. Sup-A> (enable) **set boot system flash bootflash:cat6000-sup2k8.7-2-2.bin**

In approximately 120 seconds, the image set as the boot entry on the active supervisor engine will be copied to the boot Flash on the standby supervisor engine (this is the image synchronization). This is an internal TFTP of the Cisco Catalyst OS image file and takes a few minutes to complete. The image file will have a BTSYNC appended to the beginning of the filename to designate that it has been synchronized from the active supervisor engine's boot-time image.

7. After synchronization, verify that the new image is located on the standby supervisor engine and the boot variable is properly set.

```
Sup-A> (enable) dir 2/bootflash:
```

```
Sup-A> (enable) show boot 2
```

The new Cisco Catalyst OS image is now on both supervisor engines.

8. Enable high-availability versioning on the active supervisor engine.

```
Sup-A> (enable) set system highavailability enable
```

```
Sup-A> (enable) set system highavailability versioning enable
```

You must enable versioning before the standby supervisor engine running the new software becomes active. This allows the standby supervisor engine to reboot under the new version of Cisco Catalyst OS while remaining in standby mode.

9. It is the intent of these upgrade procedures to allow for a fallback plan of using the old Cisco Catalyst OS image. The now-active supervisor engine must maintain that older image (even after an accidental reboot). Therefore, the boot variable on the active supervisor engine must be changed to its original setting, which should still be stored in the boot Flash.

```
Sup-A> (enable) set boot system flash bootflash:cat6000-sup2k8.old.bin
```

Note: Since versioning is enabled, the setting of the boot variable does not cause an image synchronization.



10. Reset the standby supervisor engine.

```
Sup-A> (enable) reset 2
```

The standby supervisor engine reboots with the new Cisco Catalyst OS image, remains in standby mode, and does not affect the operation of the active supervisor.

11. After the standby supervisor reboots, verify that it is running the new Cisco Catalyst OS image.

```
Sup-A> (enable) show module
```

The standby supervisor engine should show the new software version and it should be different from the version on the active supervisor engine.

12. Verify that the two different Cisco Catalyst OS images are high-availability compatible.

```
Sup-A> (enable) show system highavailability
```

For the high-availability switchover to occur, the operational status of the High Availability feature must be ON. If it is not, the system will be upgraded with a fast switchover (nonstateful) and the protocols will need to be restarted.

13. Reset the active supervisor engine. You will need to change the console connection to the supervisor engine in slot 2 (Sup-B) to maintain command line operation.

```
Sup-A> (enable) reset 1
```

The standby supervisor engine now takes over as the active supervisor engine (running new software) and the previously active supervisor engine is rebooted to operate in standby mode. This failover will disrupt a slight amount of traffic through the device; the amount of traffic that is affected depends on whether a high-availability switchover or a fast switchover takes place.

14. Verify the system is performing as expected. The supervisor engine in slot 2 is now active and running the new version of the Cisco Catalyst OS software. The supervisor engine in slot 1 is now in standby mode and running the previous version of software. The new standby supervisor engine can now be used as a fallback plan to revert to the previous Cisco Catalyst OS image.

15. If the system is operating as expected, the boot configuration on the standby supervisor engine (now Sup-A) will need to be updated. This can be accomplished by disabling versioning on the new active supervisor engine, which automatically enables the image synchronization feature.

```
Sup-B> (enable) set system highavailability versioning disable
```

```
Sup-B> (enable) reset 1
```

The supervisor engine Cisco Catalyst OS software upgrade procedure is now complete.

MSFC High Availability Features

The MSFC routing engine is an optional daughter card in the supervisor engine and is available in two versions, MSFC and MSFC2 (see MSFC data sheet for configuration requirements). A redundant supervisor hardware configuration can also include redundant MSFC routing engines. As such, the proper operation of the MSFC is predicated by proper operation of the supervisor engine. A supervisor reset or failover will also reset the MSFC routing engine.

The Cisco Catalyst OS High Availability feature maintains the protocol state between redundant supervisor engines, but the dual MSFCs operate under the DRM and SRM redundancy modes. Cisco recommends that you enable the Cisco Catalyst OS High Availability feature when you run either MSFC redundancy mode.



Dual Router Mode

DRM is the original MSFC configuration for redundant supervisor engine or MSFC configurations. In this mode, both MSFCs are active routers on the network. Having two active MSFCs in a single chassis does not mean having two separate routers. In fact, both MSFCs must have a nearly identical configuration, as described below in more detail. The main idea for DRM is that each MSFC independently builds an accurate picture of the Layer 3 network.

DRM Operation

The failover mechanism between MSFCs in DRM is the Hot Standby Routing Protocol (HSRP). HSRP allows the two MSFCs to maintain internal communication and react to an MSFC failover. HSRP needs to be configured on both MSFCs for each VLAN where first hop default gateway redundancy is required. Internal HSRP between MSFCs works in the same manner as HSRP between physically separate devices by sending hello messages between the routing engines. For more information about configuring HSRP, see the Cisco IOS Software Configuration Guides at:

http://www.univercd/cc/td/doc/product/software/ios121/121cgcr/ip_c/ipcprt1/1cdip.htm - xtocid26

and

http://www.univercd/cc/td/doc/product/lan/cat6000/sw_7_3/config_gd/redund.htm

Because both MSFCs have independent routing tables, there is little routing protocol convergence necessary in the event of an MSFC failure. With DRM and based on HSRP timers, the MSFC failover can be configured to less than three seconds for LAN interfaces, thus aligning the Layer 3 failover of the MSFC with the supervisor engine failover time.

Because each MSFC has the potential for taking over for the other one, they need to maintain identical configurations. This is an extremely important point to understand in DRM. Configuration parameters such as interfaces, access lists, policy routing, etc. must be configured exactly the same on both MSFCs. Parameters that cannot be duplicated on a network such as IP addresses and HSRP settings are the only parameters that are configured differently on each MSFC.

The MSFC is responsible for programming certain functions of the ASIC hardware on the PFCx. The first MSFC to go online is considered the *designated router* and the second MSFC is considered the *nondesignated router*. In a supervisor engine 1A system, both the designated router and the nondesignated router are able to program Layer 3 entries into the PFC Netflow table for routing functions. In a supervisor engine 2 system, only the designated router programs the Layer 3 entries in the PFC2s Cisco Express Forwarding table. For both a supervisor engines 1A and 2, all router ACLs and multicast shortcuts are programmed from the designated router. As you can see, the requirement for each MSFC to have an identical configuration is a necessity. If the MSFCs in DRM have different configurations, the forwarding ASICs will be programmed incorrectly resulting in unexpected behavior.

MSFC Configuration Synchronization

Beginning with the MSFC Cisco IOS Software Release 12.1(3a)E4, an MSFC redundancy feature called config-sync has been available to streamline the redundant MSFC configuration process for both MSFC and MSFC2. This feature can be used to simplify configuration of the two MSFCs and to ensure that the MSFC configurations match. Both the startup and running configurations between the designated (primary) and nondesignated (secondary) MSFCs are synchronized. Specifically, when a **write memory** or **copy <source> startup-config** command is issued on the designated MSFC, the startup configurations in NVRAM of both MSFCs are updated. This allows the configurations on the designated and nondesignated MSFCs to maintain the same configuration without having to manually type each command twice.



The following commands enable MSFC config-sync:

```
MSFC-Sup-15 (config)# redundancy  
MSFC-Sup-15 (config-r)# high-availability  
MSFC-Sup-15 (config-r-ha)# config-sync
```

With config-sync, all configurations for the designated and nondesignated MSFCs are done through the CLI of the designated MSFC. Configuration of the nondesignated MSFC is accomplished through the use of the **alt** keyword. This is the only way to configure the nondesignated MSFC when config-sync is enabled. For example:

```
MSFC-Sup-15 (config-if)# ip address a.b.c.1 x.x.x.0 alt ip address a.b.c.2 x.x.x.0  
MSFC-Sup-15 (config-if)# standby 10 priority 100 alt standby 10 priority 50
```

The command syntax does not change. The portion of the command listed before the **alt** keyword applies to the MSFC in slot 1 and the portion of the command listed after the **alt** keyword applies to the MSFC in slot 2. The config-sync feature is only supported for general IP or IPX configurations; configuration parameters for Appletalk, DECnet, etc. do not have corresponding **alt** keyword options.

WAN Interfaces in DRM

In DRM, the Optical Service Module (OSM) or FlexWAN interfaces of a WAN module are managed by only the designated MSFC. Prior to enabling the config-sync feature, the WAN interfaces do not show up in the nondesignated MSFC configuration so are not configurable on the nondesignated MSFC. During a supervisor engine or MSFC failover, the MSFC that becomes the new designated MSFC will not have properly configured WAN interfaces. For this reason a redundant supervisor or MSFC configuration without config-sync was not supported with WAN modules installed. By enabling the MSFC config-sync feature, this limitation is removed and WAN modules are supported in a redundant supervisor configuration. WAN modules should not reset during a high-availability switchover with config-sync enabled.

DRM Challenges

DRM was the original option for MSFC redundancy. This solution has been very successful by allowing for stateful Layer 3 failover between MSFCs, but it also introduces some complexity into network design and administration. The following three points present scenarios where DRM does not provide the best solution for Layer 3 redundancy:

- Each MSFC must have a unique IP address for each VLAN interface. In a distribution or core implementation using DRM as well as dual chassis, this could require up to five router IP addresses to be allocated per VLAN (four router addresses plus one HSRP address). This also increases the number of routing protocol neighbors, which can add to the CPU burden on a router. The tasks of addressing and managing four routers in this case can be a challenge that outweighs the benefits of added redundancy.
- In a redundant configuration where multiple MSFCs are connected to the same Ethernet segment, only one MSFC forwards the multicast traffic from the source to the receivers on the outgoing interfaces. The Protocol Independent Multicast designated forwarder (PIM-DF) forwards the data in the common VLAN, but the non-PIM-DF receives the forwarded multicast traffic as well. The redundant MSFC (non-PIM-DF) must drop this traffic because it has arrived on the wrong interface and will fail the reverse path forwarding (RPF) check. Traffic that fails the RPF check is called non-RPF traffic. In general, routers may not handle non-RPF traffic efficiently. With DRM, there is at least one router (the other MSFC) on each VLAN that will receive this non-RPF traffic.
- The requirement for exact configuration parameters on both MSFCs has been a complicated point for many customers. The effort to ensure that all configuration parameters are the same is a challenge when working with large Cisco IOS configuration files. Feature enhancements such as config-sync have been developed to simplify this process but do not scale.



For these scenarios, SRM is now available.

Single Router Mode

SRM is provided as an option for customers who wish to implement redundant supervisor engines or MSFCs in a system with only one active router in a chassis. SRM has the ability to use the Layer 2 and Layer 4 redundancy of Cisco Catalyst OS on the supervisor engine as well as a streamlined approach to Layer 3 redundancy. The minimum software requirements are Cisco Catalyst OS 6.3.1 and Cisco IOS Software Release 12.1(8)E2 for the MSFC.

SRM improves upon DRM. Specifically, SRM provides the following:

- A reduction in Layer 3 complexities for IP addressing and routing protocol neighbor relationships.
- A fix for the non-RPF traffic issue with having two active multicast routers on the same segment (because there is only one active router in the chassis with SRM).
- A simpler configuration for the user as only a single command set is entered from one CLI and it applies to the active router. This eliminates the challenge of ensuring that both MSFCs have the same configurations.

The following commands enable SRM:

```
MSFC-Sup-15 (config)# redundancy
MSFC-Sup-15 (config-r)# high-availability
MSFC-Sup-15 (config-r-ha)# single-router-mode
```

SRM Operation

In this mode, only the designated router will be visible to the network at any given time. The nondesignated router will be started and will maintain exactly the same configuration as the designated router (the configurations are automatically synchronized when SRM is active). However, the nondesignated router's interfaces will be kept in a line-down state and not visible to the network. Routing protocol processes are also created on the nondesignated router, but they do not send or receive updates from the network because all the interfaces are down. This is verified from the Cisco Catalyst OS command line below. Note that both the supervisor engine and the MSFC in slot 2 are listed as standby.

```
SRM> (enable) show module
Mod Slot Ports Module-Type          Model                Sub Status
-----
1   1   2   1000BaseX Supervisor      WS-X6K-SUP2-2GE     yes ok
15  1   1   Multilayer Switch Feature WS-F6K-MSFC2        no ok
2   2   2   1000BaseX Supervisor      WS-X6K-SUP2-2GE     yes standby
16  2   1   Multilayer Switch Feature WS-F6K-MSFC2        no standby
```

If the designated router fails in an SRM configuration, the other MSFC changes state from nondesignated router to designated router. This new designated router changes its interface state to link up and begins to build its routing table. It follows that the control plane failover time will be proportional to the routing protocol configuration and complexity. However, there do exist Layer 3 forwarding entries in the PFCx, which are used to forward routed traffic in the hardware path. The high availability functions of the Catalyst OS are used to maintain this forwarding information after a failover. This allows for minimal impact to the Layer 3 data plane traffic while the Layer 3 control plane converges. After the MSFC builds its routing table, the entries in the PFCx can be updated.

Beginning in Cisco Catalyst OS version 12.1(11b)E, there is a transition timer feature for running SRM on the supervisor engine 2/PFC². This timer configures the time that the new designated router will wait before downloading any new hardware Cisco Express Forwarding entries to the PFC2. Due to differences in routing convergence times, the default of 120 seconds might not be long enough to allow for complete convergence before programming the PFC2 hardware.



The same IP and Media Access Control (MAC) addresses are used for the designated router, whether or not the MSFC is the designated router. The MSFC that is chosen as the designated router will communicate its default MAC address to the MSFC that is the nondesignated router. All subsequent interfaces created on the nondesignated router use this MAC address, unless the user explicitly configures a different MAC address.

On bootup the two MSFCs perform a “handshake” process, which takes about a minute, before entering SRM mode.

Important: Do not make configuration changes on the nondesignated router during the handshake process.

The following commands can be used to verify that SRM is enabled:

```
SRM# show redundancy
Designated Router: 1 Non-designated Router: 2
Redundancy Status: designated
Config Sync AdminStatus : enabled
Config Sync RuntimeStatus: enabled
Single Router Mode AdminStatus : enabled
Single Router Mode RuntimeStatus: enabled
Single Router Mode transition timer : 120 seconds
```

For more details about configuring SRM, see section “MSFC Redundancy–Single Router Mode Redundancy” in the Cisco Catalyst OS configuration guide at:

http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/sw_6_3/config_gd/redund.htm

WAN Interfaces in SRM

Because the MSFC configurations are synchronized as an inherent part of SRM, all OSMs and FlexWAN WAN modules are supported with redundant supervisor engines or MSFCs configured for SRM. As in DRM, the designated router manages the WAN interfaces. The interfaces are configured completely on the designated router and that configuration is synchronized to the nondesignated router. For failover scenarios, the new designated router will take ownership of the WAN interfaces as soon as that MSFC becomes the designated router. Additionally, the WAN modules should not reload upon a high-availability switchover. With SRM enabled, there is no manual configuration necessary on the WAN interfaces to support an MSFC failover.

SRM Configuration and Conversion Procedure

The configuration guide has a very good procedure for configuring SRM, converting from DRM to SRM, and performing software upgrades with SRM enabled. The latest recommended procedures are available at the following URLs:

“Configuring Single Router Mode Redundancy”

http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/sw_7_3/config_gd/redund.htm#1071789

“Upgrading Images with Single Router Mode Enabled”

http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/sw_7_3/config_gd/redund.htm#1071832

2. This feature does not apply to the supervisor1a/PFC/MFSCx because the PFC uses flow-based forwarding architecture and all new flows are initially sent to the MFSCx software path.



SRM and IP Multicast

With SRM, the VLAN interfaces on the nondesignated router are in a *down* state. Even after a failover, these interfaces will not move into an *up* state until the supervisor engine verifies that the VLAN has at least one connected physical interface in the forwarding state. This interruption causes the supervisor engine to delete all multicast entries in the PFCx and this disrupts multicast forwarding. As an enhancement to the original SRM implementation, Cisco Catalyst OS release 7.1 provides support for IP Multicast stateful redundancy. When SRM is enabled in Cisco Catalyst OS version 7.1, the multicast flows are preserved during a failover.

Supervisor and MSFC Failover Tests

A test environment was created to demonstrate sample high-availability switchover scenarios and record the corresponding failover times. The test setup included one Cisco Catalyst 6509 chassis with dual supervisor engine 2 line cards running Cisco Catalyst OS version 7.2.2 or MSFC2 hardware running Cisco IOS Software Release 12.1(11b)E4. These tests were intended to be basic and easily demonstrated. The testing mechanism was pings between two end devices connected directly to the switch. Spanning-Tree was enabled on the ports in use for all scenarios. Tests were initiated by resetting the switch supervisor engine. Each scenario was tested eight times and the results were averaged.

Layer 2 Failover

For Layer 2 traffic between two end stations within the same VLAN, the failover resulted in a one or two ping timeout (approximately one- or two-second failover time).

Note: This test can be performed on a supervisor engine without an MSFC.

Layer 3 Failover

For Layer 3 traffic, a common test environment was created for measuring the supervisor or MSFC failover of a single Cisco Catalyst 6500 Series configured first with DRM and then with SRM. The two pinging devices were placed in separate VLANs. The basic software configuration included enabling the High Availability feature on the Cisco Catalyst OS and then either DRM or SRM redundancy in the Cisco IOS Software on the MSFC2. The complete configurations are shown below.

SRM

```
hostname SRM
!
redundancy
  high-availability
  single-router-mode
!
interface Vlan20
  ip address 10.20.1.3 255.255.255.0
  no ip redirects
!
interface Vlan30
  ip address 10.30.1.3 255.255.255.0
  no ip redirects
!
end
```




DRM

```
hostname DRM
!
redundancy
  high-availability
  config-sync
!
interface Vlan20
  ip address 10.20.1.3 255.255.255.0 alt ip address 10.20.1.2 255.255.255.0
  standby ip 10.30.1.4
  standby priority 100 alt standby priority 50
  no ip redirects
!
interface Vlan30
  ip address 10.30.1.3 255.255.255.0 alt ip address 10.30.1.2 255.255.255.0
  standby ip 10.30.1.4
  standby priority 100 alt standby priority 50
  no ip redirects
!
end
```

For DRM, the average failover time was measured at 2.56 seconds. For SRM, the average failover time was measured at 2.31 seconds. Keep in mind that both DRM and SRM will maintain Layer 3 forwarding entries in the hardware-forwarding table (a function of the Cisco Catalyst OS High Availability feature). The difference is that DRM employs two active routers in the chassis and the SRM employs only one router. So SRM requires a routing table recalculation in the software that DRM does not, but neither has a direct effect on the failover time for Layer 3 traffic.

A second test with one workstation running a File Transfer Protocol (FTP) client and another running an FTP server was run. Transferring a 10-MB file across the switch at Layer 3 during normal operations took an average of 16 seconds. If a supervisor switchover takes place during this same FTP session, the transfer time is an average of 18 seconds. During switchover, the difference in FTP transfer time is only 2 seconds. This example is used to demonstrate a realistic TCP application.

A third test with IP phones connected to the Cisco Catalyst 6500 Series involved establishing an IP phone call between one local IP phone and one remote IP phone. With the Cisco Catalyst OS High Availability feature enabled, a supervisor engine switchover was initiated. The IP phone call was maintained through the switchover and the call participants noticed minimal disruption. This provides a real-world example of the capability of the Cisco Catalyst 6500 to provide high availability in all layers of the network.

For general purposes, it is still maintained that the Layer 2 and Layer 3 stateful supervisor engine switchover will take place in less than three seconds to cover most real-world scenarios.

Redundant Power Supplies

Currently the Cisco Catalyst 6500 Series can be configured with a 1000- (6-slot chassis only), 1300-, 2500-, or 4000-watt power supply. In addition, a power redundancy feature is available to further specify the power configuration via software. By default, the power redundancy (or load sharing) feature is enabled. With two power supplies installed and power redundancy enabled, the total power drawn from both supplies is at no time greater than the capability of one supply. If one power supply fails, the other power supply can take over the load and the complete system remains powered. If power redundancy is



disabled, the power available to the system is the combined power of the power supplies. Note that in this configuration, if one supply fails, the system might not have enough power to supply all the modules. To enable or disable power redundancy, the following command is used:

```
Sup-A> (enable) set power redundancy enable | disable
```

Each line card has different power requirements so power requirements vary from chassis to chassis. The user guide documents the Cisco Catalyst 6500 Series power requirements to help you understand the power requirements for each particular line card. These requirements are described at:

http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/6000hw/inst_aug/02prep.htm

Conclusion

The High Availability and Redundancy features of the Cisco Catalyst 6500 Series provide a very reliable switching and routing platform. The hardware redundancy provided by dual supervisor engines, dual routing engines, dual switching fabrics, and a dual power supply help to reduce the potential downtime of a network. Software redundancy features such as the High Availability feature of Cisco Catalyst OS and the DRM or SRM options for MSFC failover build on this hardware redundancy to create a very stable operating environment. The combination of these features, in addition to the system performance and intelligent network services, makes the Cisco Catalyst 6500 Series unparalleled in the industry.



Corporate Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

www.cisco.com

Tel: 408 526-4000

800 553-NETS (6387)

Fax: 408 526-4100

European Headquarters

Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands

www-europe.cisco.com

Tel: 31 0 20 357 1000

Fax: 31 0 20 357 1100

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

www.cisco.com

Tel: 408 526-7660

Fax: 408 527-0883

Asia Pacific Headquarters

Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912

www.cisco.com

Tel: +65 317 7777

Fax: +65 317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the Cisco Web site at www.cisco.com/go/offices

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

All contents are Copyright © 1992–2002, Cisco Systems, Inc. All rights reserved. Catalyst, Cisco, Cisco IOS, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0208R) MH/LW3733 12/02